Supporting Information For:

Quantifying landscape change following catastrophic dam failures in Edenville and Sanford, Michigan, USA

Harrison K. Martin[1,2,3]*, Douglas A. Edmonds[1], Brian J. Yanites[1], and Nathan A. Niemi[4]

[1]*Department of Earth & Atmospheric Sciences, Indiana University, 1001 East 10th Street, Bloomington, IN 47405-1405*
[2]*Division of Geological and Planetary Sciences, California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125*
[3]*Resnick Sustainability Institute, California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125*
[4]*Department of Earth and Environmental Sciences, University of Michigan, 1100 North University Avenue, Ann Arbor, MI 48109-1005*
**Corresponding author: hkm@caltech.edu*

DOI: 10.1002/esp.5855

1. Aerial lidar methodology:

The USGS 3DEP lidar data set is available as two separate point clouds, one for Midland County (2017):

[https://portal.opentopography.org/usgsDataset?dsid=USGS_LPC_MI_31Co_Midland_2016_LAS_2019]

and another for Gladden County (2016-2017):

[https://portal.opentopography.org/usgsDataset?dsid=USGS_LPC_MI_31Co_Gladwin_2016_LAS_2019].

We merged the two point clouds, unifying projections. Point densities were 2.94 pt/m^2 for the Edenville inundation area for 2.38 pt/m^2 for Sanford. The Sanford point cloud was laterally shifted by 0.44 m east and 0.63 m south to correct for a global offset and match the UAV-based lidar point clouds. The source point clouds were originally provided with projections of State Plane Michigan South and State Plane Michigan Central, respectively, with vertical units of international feet. All

spatial data, including UAV-based lidar data, were unified with horizontal and vertical units of meters and reprojected to NAD83(2011) and NAVD88.

2. UAV-based lidar methodology

2.1 Summary/Overview (details follow in sections below):

We collected three UAV-based lidar datasets at the Edenville and Sanford sites. We used stationary GPS stations to apply differential corrections to the UAV flightpaths. We fused lidar sensor data to positional data to create point clouds, which we then classified in order to look only at bare-earth points. We applied corrections to the points within individual point clouds to reduce between-flightline errors. We measured the distance between points from different point clouds (collected on different dates) in control, fiducial areas that were believed to have not experienced topographic change (such as roads and parking lots) and applied global, constant shifts to correct these observed differences in order to align point clouds and reduce error. The range of mean differences across fiducial areas was used as an estimate of potential mean error. Finally, for changes downstream of the dam failures, we measured the distance between aligned point clouds (collected on different dates) and calculated the 2.5D volume of these differences.

2.2 Data collection:

UAV-based lidar data were collected using a Phoenix LiDAR Systems miniRANGER mounted on a DJI M600 with onboard GNSS/INS, including a gyroscopic IMU, dual-GPS, and differential real-time kinematics. The miniRANGER has a 905 nm laser with a 100 kHz scan rate and up to five returns per pulse, allowing for vegetation penetration. Collections at both Edenville and Sanford were performed May 31-June 2, 2020, August 07-09, 2020, and April 21, 2021. Edenville necessitated three base stations in series based on different UAV launch points, while Sanford needed only one. Point densities for all UAV-based lidar scans ranged from 7.72 to 9.41

53 pt/m^2 after ground classification. At each site, we used a CHC X900 GNSS to collect >2 hours of
54 stationary data, to allow for correction via the NOAA Online Positioning User Service, yielding
55 an inherent, global-shift root-mean-squared-error (RMSE) of approximately 1 cm (Supplemental
56 Table 1).

57 2.3 Lidar processing:

58 We used these base stations to apply differential corrections to the UAV rover's positions
59 via NovAtel Inertial Explorer software. We then fused lidar data to trajectories to yield point clouds
60 using proprietary Phoenix LiDAR Systems Spatial Explorer v.4 software. We used the proprietary
61 TerraScan and TerraMatch software and algorithms, by TerraSolid, to apply ground classification
62 to data [https://terrasolid.com/guides/tscan/crground.html]. We selected a maximum building size
63 of 60 m, a maximum terrain angle of 88 degrees, an iteration angle of 6 degrees to plane, an
64 iteration distance of 1.4 meters to plane, and opted to reduce iteration angle when edge lengths
65 were less than 5 meters. We then applied a series of global, flightline, and fluctuating corrections
66 that minimized the RMSE of tie points observed between different flightlines. This yielded a
67 measured between-flightline average error magnitude $\leq$ 4 cm (Supplemental Table 1). This
68 average error magnitude is greater than the average error, as errors are approximately evenly
69 distributed about the mean.
70

Supporting Table 1: GNSS base station uncertainties and within-point-cloud errors (based on tie points between flights) for each site. All measurements are in meters and rounded to three decimal places.

Location	Date	OPUS Overall Base Station RMSE				Microstation Average Magnitude	Initial Registration Error (sum)
		Bacon	Consumer	Causeway	Sanford		
Edenville	2020-06	0.013	0.014	0.013	--	0.039	0.079
	2020-08	0.014	0.011	--	--	0.028	0.054
	2021-04	0.011	--	--	--	0.040	0.051
Sanford	2020-06	--	--	--	0.012	0.040	0.052
	2020-08	--	--	--	0.016	0.031	0.047
	2021-04	--	--	--	0.012	0.035	0.047

2.4 Data Alignment:

Data were imported to CloudCompare [https://www.danielgm.net/cc/]. We further refined the accuracy of our comparisons by correcting for global vertical offsets that exist between datasets. To test for and determine the magnitude of these offsets, we first compared the distance between point clouds using the M3C2 plug-in. Scales were set to 2.5 m for normal, 2.5 m for projection, and 10 m for depth. We used every point in the earlier cloud as the core points. For normals, we used the multi-scale difference option, with a minimum normal of 0.25 m, a step of 0.25 m, a maximum of 10 m, and a preferred orientation of +Z. For only this initial comparison, registration error was set to the value in the column “Initial Registration Error” in Supplemental Table 1. This value is equal to the sum of the OPUS-reported Overall RMSE (m) for the base stations for each mission within the point cloud, plus the average magnitude of error reported

within Microstation for each point cloud, after global, flightline, and fluctuating corrections have been applied (Supplemental Table 1).

We then established fiducial areas. These areas are regions in which, through satellite imagery, lidar, field observations, and UAV imagery, we do not believe true geomorphic change occurred between the two dates of observation. If this belief is true, then any change observed in these areas between observations must be due to error. If these errors are roughly consistent across the study area, a linear, global offset can be used to align point clouds and minimize overall errors. Good candidate regions include roadways, parking lots, and similar flat anthropogenic areas. Within each area, we fit a Gaussian curve to the histogram of all point-to-point offsets. The mean and standard deviation of this curve was recorded. We interpret the mean as the bias or the offset. The standard deviation, instead, reflects the precision of the measurement and it is a function of how consistently and precisely each individual surface is measured; the mean and standard deviation of this curve reflect accuracy and precision, respectively. To ensure that the areas were 1) genuinely flat and 2) undisturbed between scans, we measured the standard deviations of i) the elevation distribution of points within each individual scan, and ii) the distribution of differences between the scans, for each fiducial area. We looked for values that were small and consistent, respectively. We attempted to space these fiducial areas out across the study area so as to not minimize error in any one small area at the expense of the overall dataset. We only accepted areas that had hundreds or thousands of point-to-point comparisons. Most fiducial areas within each study area showed the same directionality of change, with similar magnitudes and standard deviations and approximately normal distributions about their means (Supplemental Tables 2-7). Finally, we took the calculated mean differences, weighted by the number of observations in each fiducial area, and applied them globally to the later of the two flights involved in each difference.

Repeating the M3C2 process on the aligned data yields mean differences in fiducial areas that appear normally distributed, with similar standard deviations as before. The mean differences across fiducial areas are not identical; for the 3DEP to 2020-06 differences as an example, they vary by 0.033 m in Edenville (-0.018 to -0.051m) and 0.025 m in Sanford (0.075-0.100 m). This remaining spread in fiducial area offsets indicates that the systematic offset is not perfectly uniform, and reflects measurements of irreducible errors. As a result, and because these errors observed in fiducial areas already incorporate all misregistration errors, we use these spreads of mean differences across fiducial areas as our estimate of potential mean area for ± values in volumetric calculations. One limitation of this study worth noting is that all fiducial areas lack vegetation and are flat, which means that they lack errors associated with vegetation penetration, ground classification, or steep slopes. As there were no fiducial areas in vegetated or uneven terrain, they likely underestimate error when applied to vegetated or steep terrain in our study area.

Supporting Table 2: Fiducial areas and offsets observed between the 2016 3DEP aerial lidar and 2020-06 UAV-based lidar point clouds for Edenville. Locations are in Supplemental Figure 1. All distances are in meters and are rounded to three or six decimal places. Negative values are erosion.

fiducial area name	# of compared points in area	mean difference	standard deviation	contribution to total weighted mean difference
North Causeway-2	29,614	-0.031	0.027	-0.023096
South Causeway-2	6,815	-0.030	0.028	-0.008683
Mid Xsec East	1,969	-0.051	0.018	-0.003996
Mid Highway East	3,633	-0.044	0.022	-0.001753
SE Highway	1,851	-0.021	0.022	-0.001176
Mid Highway West	3,877	-0.022	0.027	-0.001838
SW Road	2,287	-0.018	0.019	-0.000442
Sums:	**50,046**	--	--	**-0.030**

Supporting Table 3: Fiducial areas and offsets observed between the 2020-06 and 2020-08 UAV-based lidar point clouds for Edenville. Locations are in Supplemental Figure 3. All distances are in meters and are rounded to three or six decimal places. Negative values are erosion.

fiducial area name	**# of compared points in area**	**mean difference**	**standard deviation**	**contribution to total weighted mean difference**
North Causeway-3	4215	0.007	0.005	0.001
South Causeway-2	7003	0.009	0.012	0.003
Mid Highway East-2	2514	-0.010	0.005	-0.001
SE Highway-2	2042	-0.002	0.005	-0.000
Mid Highway West	3917	-0.006	0.009	-0.001
Sums:	**19691**	**--**	**--**	**0.002**

Supporting Table 4: Fiducial areas and offsets observed between the 2020-08 and 2021-04 UAV-based lidar point clouds for Edenville. Locations are in Supplemental Figure 5. All distances are in meters and are rounded to three or six decimal places. Negative values are erosion.

fiducial area name	**# of compared points in area**	**mean difference**	**standard deviation**	**contribution to total weighted mean difference**
Lower Highway East	2655	-0.001	0.008	-0.000
SE Highway-2	2255	-0.008	0.005	-0.002
Mid Highway West	5644	-0.005	0.013	-0.003
Sums:	**10554**	**--**	**--**	**-0.005**

Supporting Table 5: Fiducial areas and offsets observed between the 2016 3DEP aerial lidar and 2020-06 UAV-based lidar point clouds for Sanford. Locations are in Supplemental Figure 6. All distances are in meters and are rounded to three or six decimal places.

fiducial area name	**# of compared points in area**	**mean difference**	**standard deviation**	**contribution to total weighted mean difference**
Center Intersection	1,516	0.075	0.013	0.011223
West Intersection	1,100	0.080	0.015	0.008675
East Intersection	1,294	0.094	0.016	0.011998
NW Parking Lot	478	0.082	0.040	0.003862
NE Roads	5,716	0.100	0.024	0.056845
Sums:	**10,104**	**--**	**--**	**0.093**

Supporting Table 6: Fiducial areas and offsets observed between the 2020-06 and 2020-08 UAV-based lidar point clouds for Sanford. Locations are in Supplemental Figure 8. All distances are in meters and are rounded to three or six decimal places.

fiducial area name	**# of compared points in area**	**mean difference**	**standard deviation**	**contribution to total weighted mean difference**
Center Intersection	1,514	0.022	0.008	0.00621
West Bridge	1,483	-0.002	0.006	-0.000650
NW Parking Circle	678	-0.007	0.007	-0.001010
South Road	1120	0.028	0.009	0.006586
Sums:	**4,795**	**--**	**--**	**0.012**

Supporting Table 7: Fiducial areas and offsets observed between the 2020-08 and 2021-04 UAV-based lidar point clouds for Sanford. Locations are in Supplemental Figure 10. All distances are in meters and are rounded to three or six decimal places.

fiducial area name	# of compared points in area	mean difference	standard deviation	contribution to total weighted mean difference
SSW Road	772	-0.009	0.010	-0.002790
Central Road	307	-0.001	0.012	-0.000150
NW Parking Circle	1377	-0.020	0.008	-0.011300
Sums:	**2,456**	**--**	**--**	**-0.014**

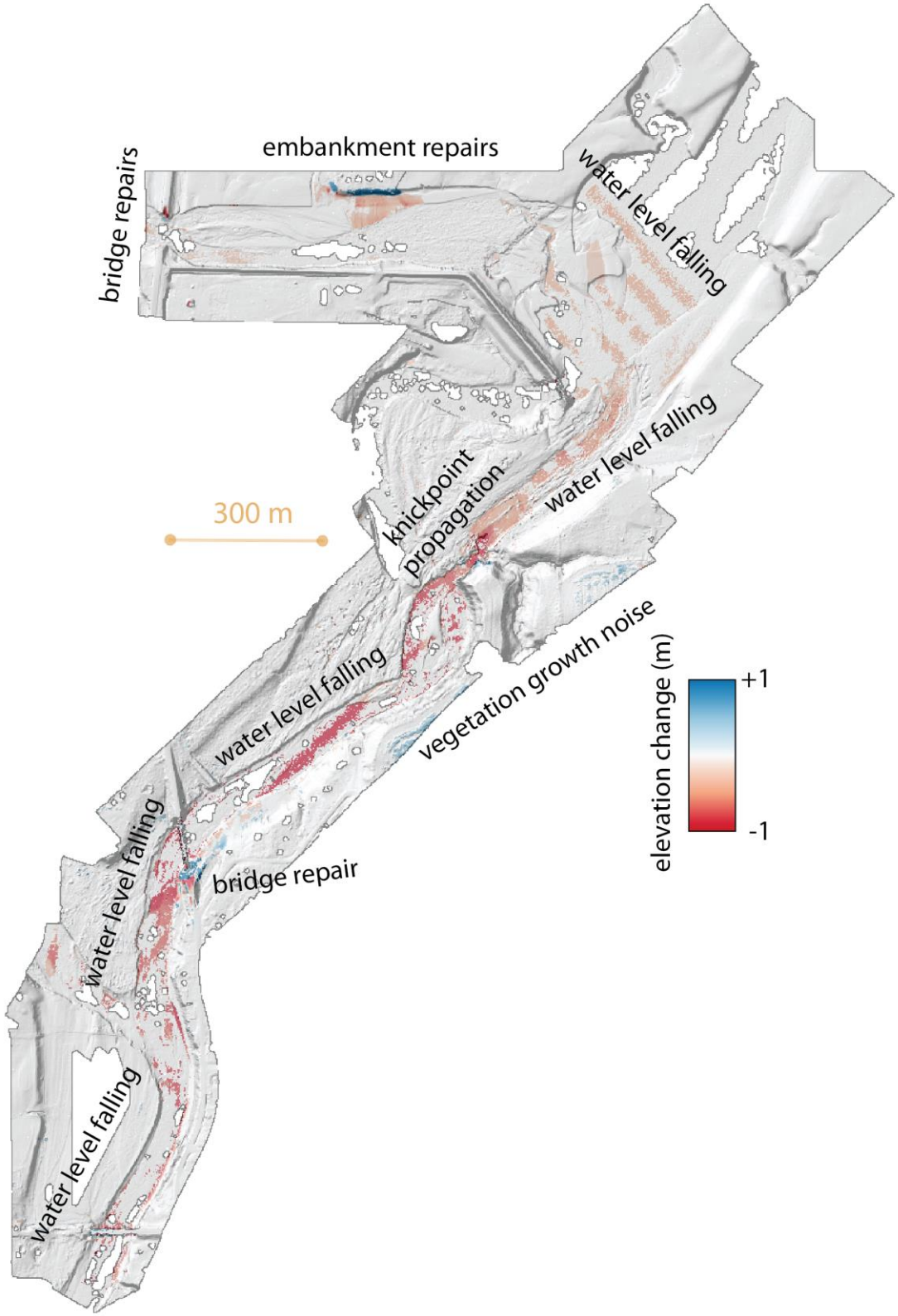
2.5 Volumetric calculations:

Once we have applied global offset corrections to the latter point cloud in each difference pair, we can once again calculate M3C2 distances using CloudCompare. We use the same settings as above, except that we can now use the registration errors that were estimated at the end of Supplemental Section 2.4: for 3DEP to 2020-06, these are 0.033 m for Edenville and 0.025 m for Sanford. The M3C2 algorithm calculates spatially variable confidence intervals based on these registration errors and local point cloud roughnesses as measured along the local normal directions. The algorithm uses this to classify whether observed changes are statistically significant at a 95% confidence interval level of detection. We discard all points that do not have significant changes. Taking this point cloud of only significant differences, we discard all points that lie outside the inundation and data extents shown in Figure 2. Next, we assign the M3C2 distance to the Z-coordinate of each point. We then calculate the 2.5D volume of this new point cloud of significant differences by leaving empty cells empty and specifying a 2 m rasterization size, with volumes calculated along the Z axis and the value of each cell equal to the average height of all points

contained within it. The 2 m rasterization effectively fills in small gaps; the largest gap that could be theoretically filled in is 5.66 m. Calculating the 2.5D volume reports the volume eroded, the volume deposited, the net change, and the total area (along the XY plane) of the area that experienced significant changes. We apply this to the study areas downstream of the dam failures as marked in Figure 2, and each sub-region as marked in Figure 6. For each region, we calculate mean vertical changes by dividing volume changes by areas that experienced change (including both erosion and deposition). We calculate ± error bounds by multiplying the registration error for each site (as calculated at the end of Supplemental Section 2.4) by the area of significant change (including both erosion and deposition) at each site. Changes after the 2020-06 flood are minor and attributable to, noise, changes in water level, or anthropogenic changes; these are qualitatively presented in Supplemental Figures 1-10, but are otherwise not quantitatively presented.



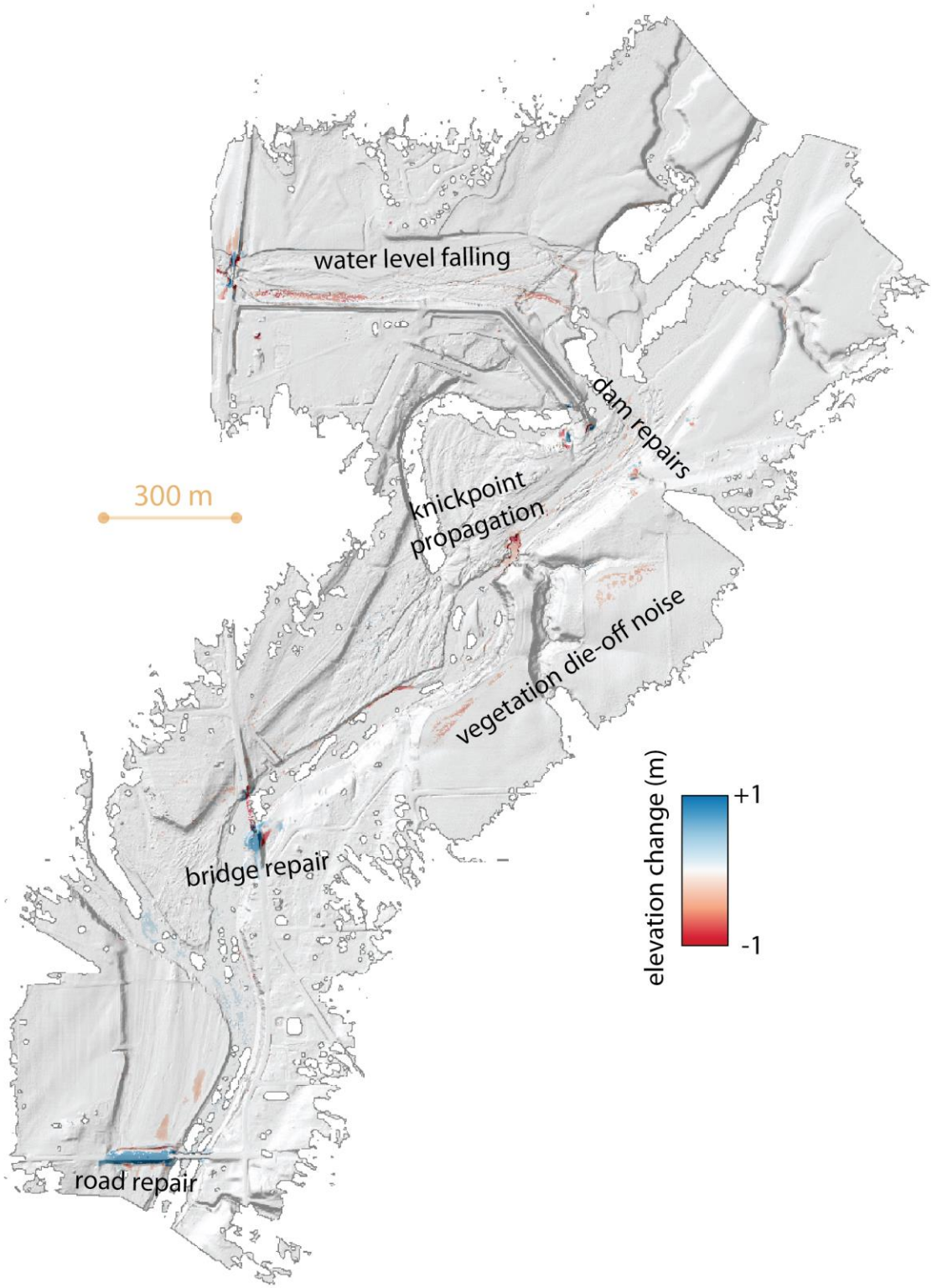
Supporting Figure 1: Hillshaded topography of the Edenville site from UAV-based lidar on May 31-June 2, 2020. Fiducial areas traced in white and labelled. North is up for Supplemental Figures 1-10.



Supporting Figure 2: M3C2 changes in Edenville between UAV-based lidar scans on May 31-June 2, 2020 and August 7-9, 2020, overlain atop a hillshade of topography from the latter scan.

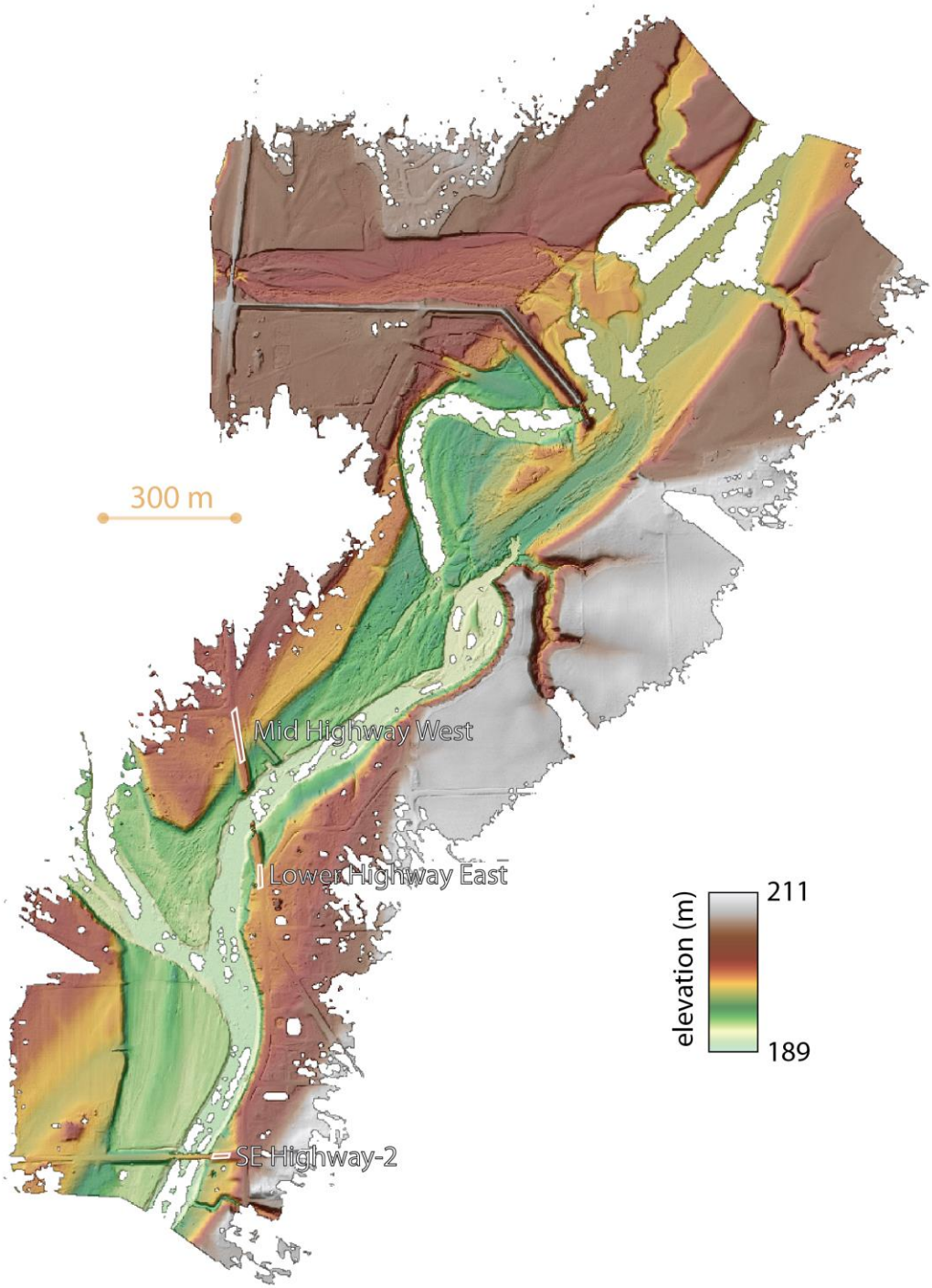


Supporting Figure 3: Hillshaded topography of the Edenville site from UAV-based lidar on August 7-9, 2020. Fiducial areas traced in white and labelled.



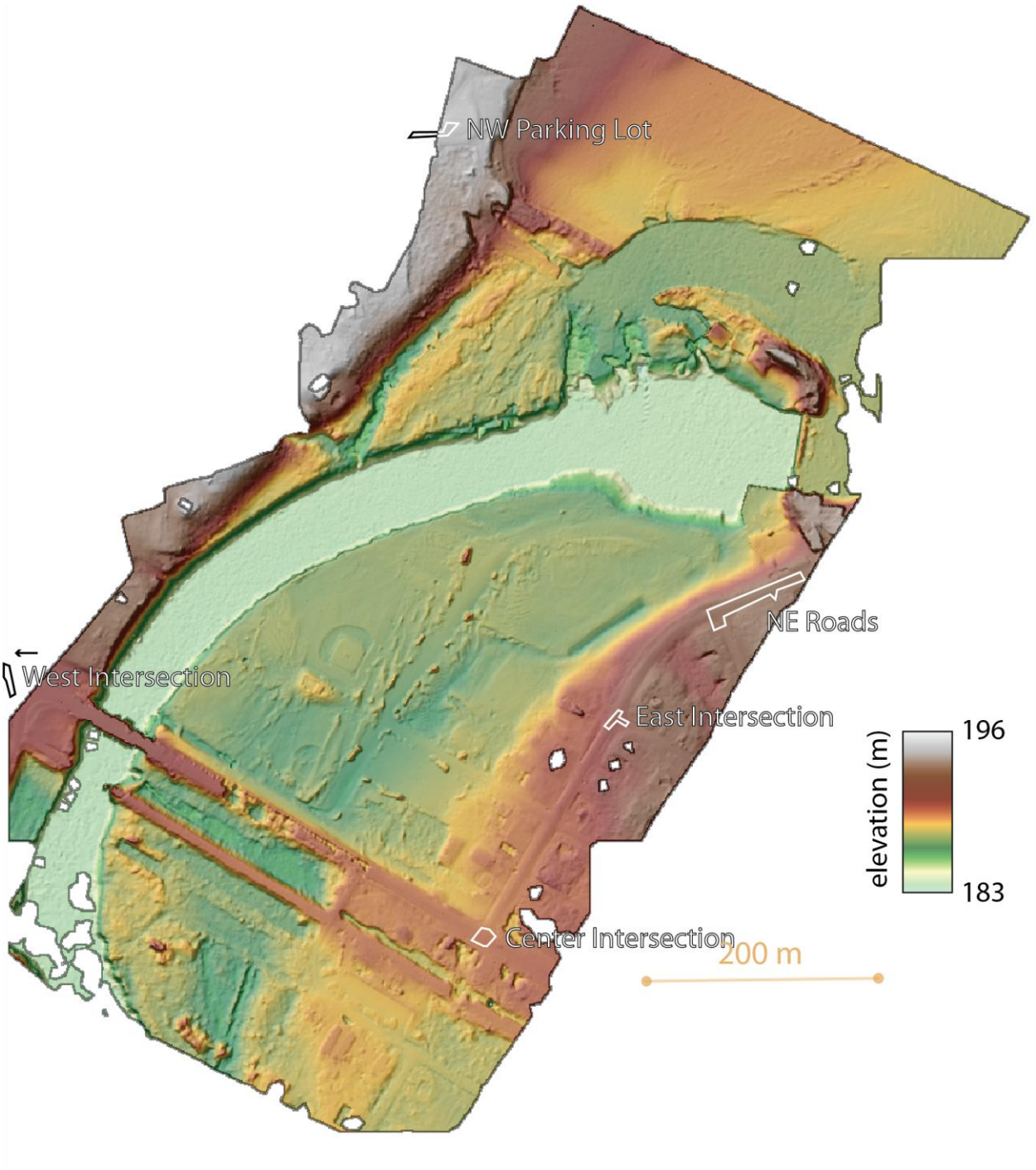
Supporting Figure 4: M3C2 changes in Edenville between UAV-based lidar scans on August 7-9, 2020 and April 21, 2021, overlain atop a hillshade of topography from the latter scan.

188

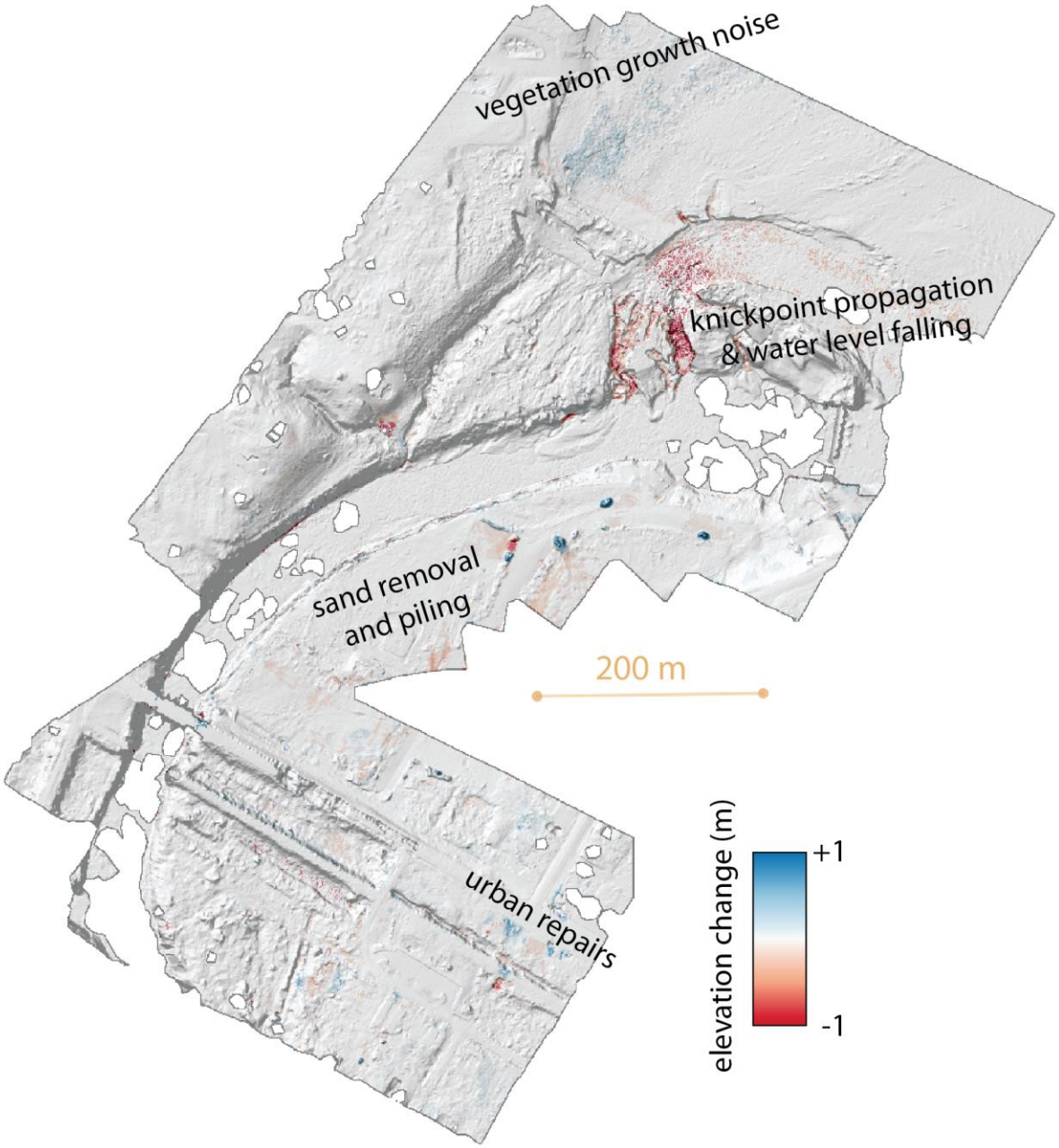


189

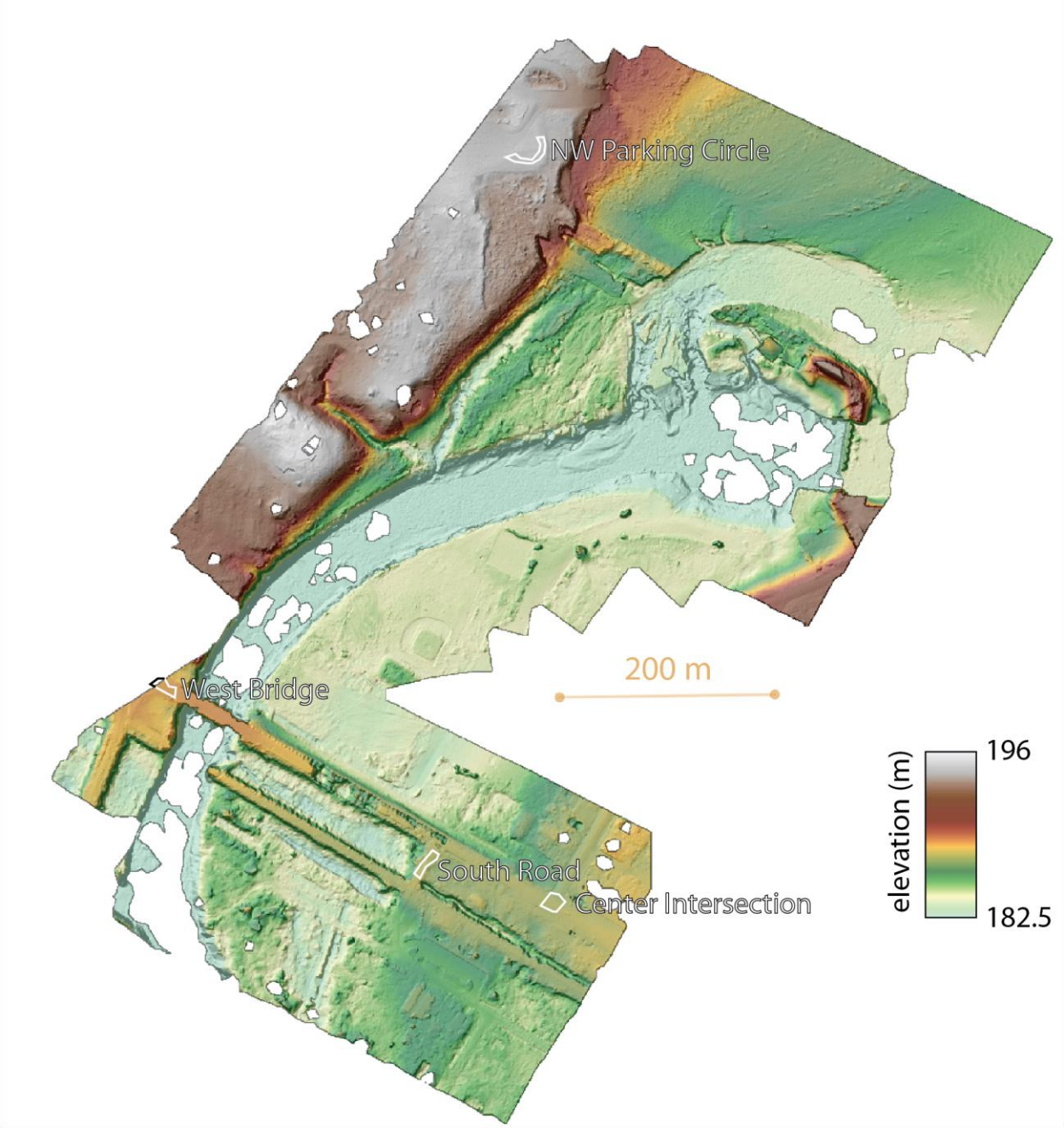
190 **Supporting Figure 5: Hillshaded topography of the Edenville site from UAV-based lidar on**
191 **April 21, 2021. Fiducial areas traced in white and labelled.**



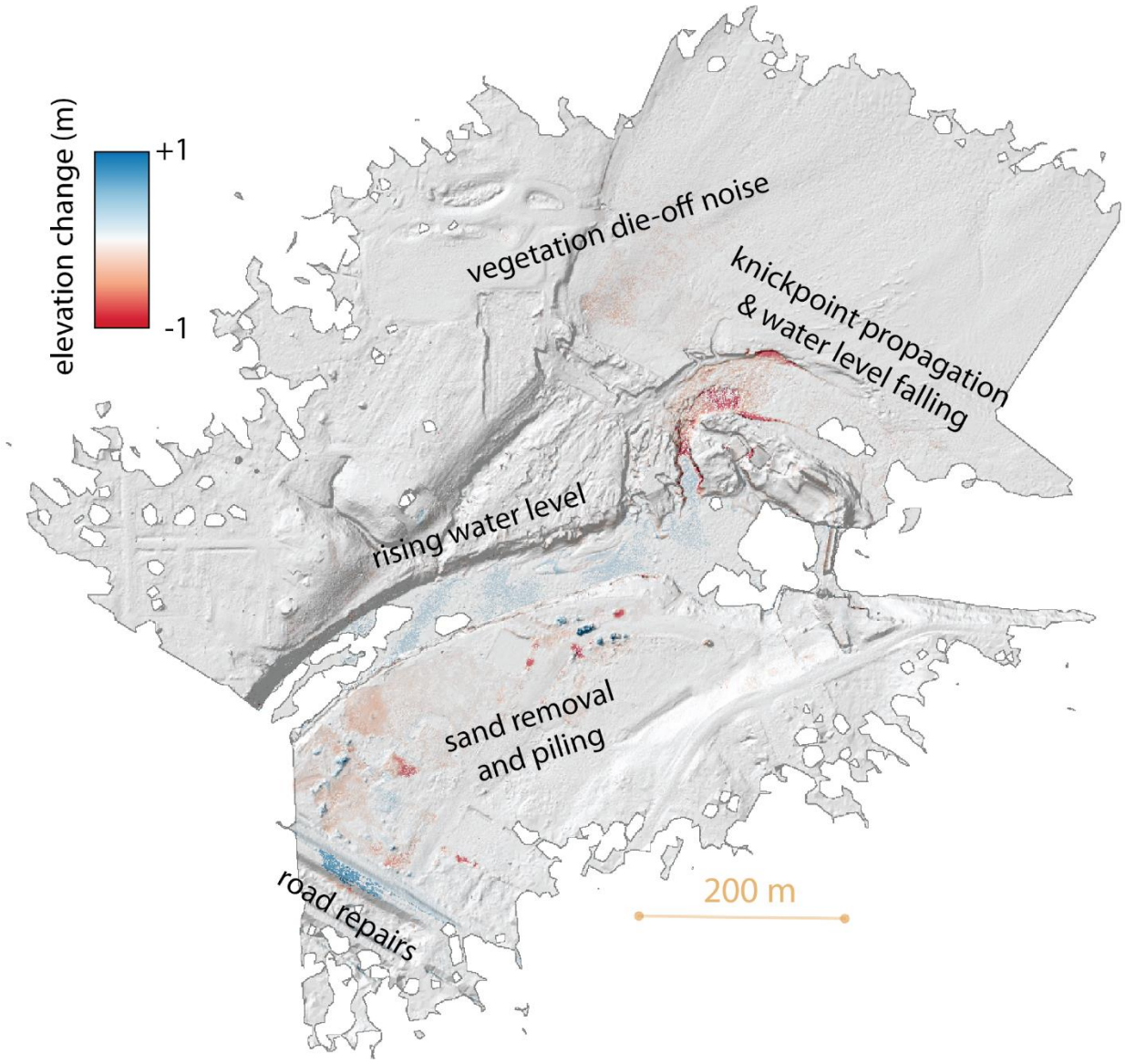
Supporting Figure 6: Hillshaded topography of the Sanford site from UAV-based lidar on May 31-June 2, 2020. Fiducial areas traced in white and labelled.



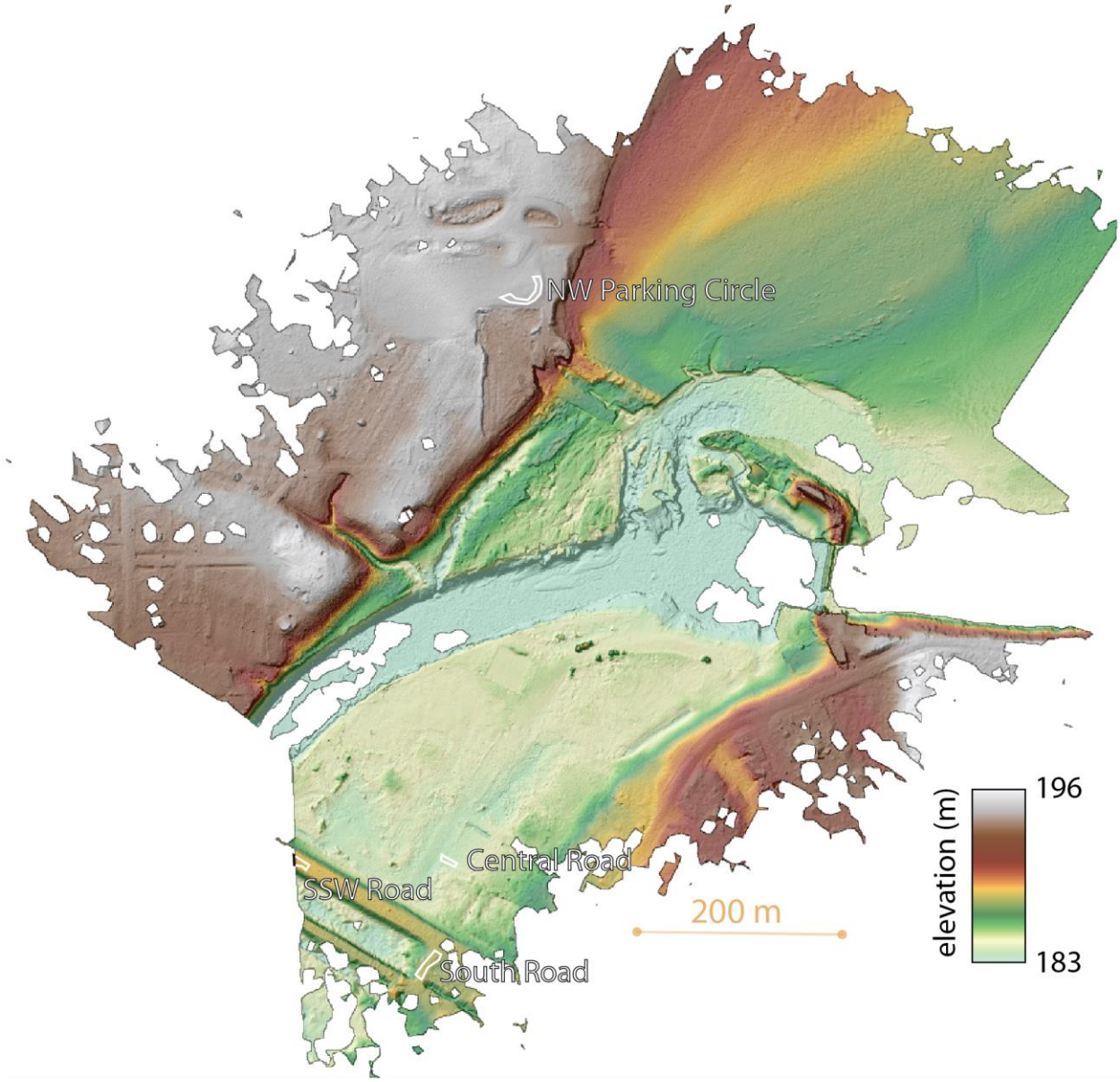
Supporting Figure 7: M3C2 changes in Sanford between UAV-based lidar scans on May 31-June 2, 2020 and August 7-9, 2020, overlain atop a hillshade of topography from the latter scan.



Supporting Figure 8: Hillshaded topography of the Sanford site from UAV-based lidar on August 7-9, 2020. Fiducial areas traced in white and labelled.



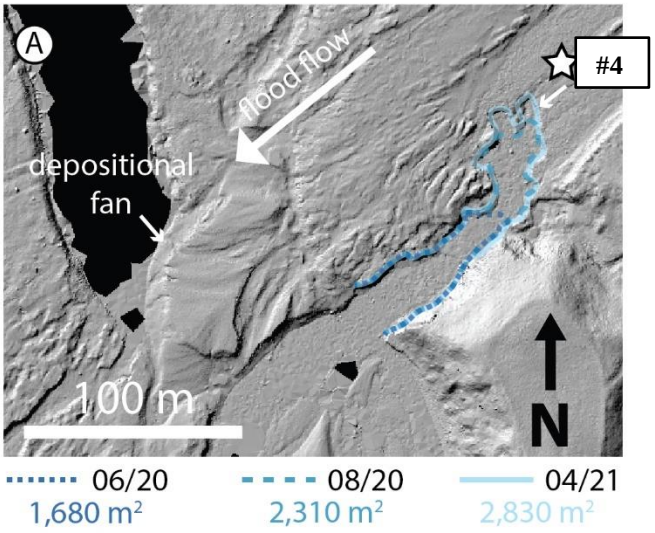
Supporting Figure 9: M3C2 changes in Sanford between UAV-based lidar scans on August 7-9, 2020 and April 21, 2021, overlain atop a hillshade of topography from the latter scan.



Supporting Figure 10: Hillshaded topography of the Sanford site from UAV-based lidar on April 21, 2021. Fiducial areas traced in white and labelled.

2.6 Knickpoint propagation:

Knickpoint propagation was measured on three dates: May 31-June 2, 2020, August 7-9, 2020, and April 21, 2021 (Supplemental Figures 11-13). Propagation was measured by tracing the scarp edge from DEMs derived from UAV-based lidar surveys. As such, propagation was only measured along the XY plane and has no volumetric meaning. Scarp retreat on the upstream end of area (iv) was also imaged (Supplemental Figures 14-15).



Supporting Figure 11: Detailed view of the propagation of Knickpoint #4 overlain on a hillshaded DEM collected on April 21, 2021.



Supporting Figure 12: Field photo taken on 8 August 2020, looking NW while standing on the east side of Knickpoint #4. Note the poorly sorted but unstratified material that makes up the knickpoint. Despite months and tens of meters of propagation, the edge of the knickpoint remains sharp with a nearly vertical erosional face.



Supporting Figure 13: Field photo taken on 2 April 2021, looking NW while standing on the east side of Knickpoint #4. Despite the knickpoint propagating a further ~13 m upstream over the nearly eight months since Supplemental Figure 12, the edge of the knickpoint remains sharp with a nearly vertical face.



Supporting Figure 14: Google Earth satellite image from 15 June 2020 showing the confluence of the Tobacco and Tittabawassee Rivers downstream of the Edenville Dam post-failure. White line shows pre-failure bankline. Red star marks the position of the photographer in Supplemental Figure 15, looking ESE. Yellow scalebar is 10 m. Center of image is 43.804, -84.389.



Supporting Figure 15: Field photo showing more than 10 m of local bank erosion due to dam failure at the Edenville site. Flood flow direction was to the right. Position of photographer shown in Supplemental Figure 14, looking ESE.

3. Flood hydrology methodology

We estimated the peak discharge of both the Edenville and Sanford Dam failures following Walder and O'Connor (1997). Values for all parameters are given in Supplemental Table 8. Those authors use a physically based model of breach formation and lake drainage and data from past breaches to argue that dimensionless peak discharges can be plausibly estimated from common and simply measured variables. We followed the suggested method for predicting peak discharge at the breach in section 6 on page 2345. Following their steps for each of our two dam failures:

1. The authors first suggest trying to calculate a dimensionless variable η, which is the product of dimensionless lake volume and dimensionless vertical breach erosion rates. We calculated η using the definition equation on page 2342,

$$\eta = \frac{k * V_0}{g^{\frac{1}{2}} * d^{\frac{7}{2}}}$$ (Supplemental Eq. 1)

where k is the mean vertical erosion rate of the breach (m/s; itself calculated as $\frac{d}{t}$ where d is the drop in water level or the thickness of breach eroded [m] and t is the breach-formation time [s]), V_o is the volume of water released, and g is acceleration due to gravity (9.81; m/s^2).

   - d: measured from post-failure lidar as the elevation drop at the edge of the remaining embankment adjacent to the failure. Our values fall within the typical range of 50-100% of pre-failure dam height reported by Walder and O'Connor (1997).
   - t: France et al. (2022) report that, for the Edenville Dam, "the embankment was fully breached" in less than 10-30 seconds from when failure began. While they also report that "the breach enlarged over the next few hours, releasing the water stored in Wixom Lake", this enlargement seems to be mostly due to widening. We use an estimate of 1 minute for breach failure. Note that this is much quicker than for many other dam breaches (Walder and O'Connor 1997), as this failure occurred due to static liquefaction. The Sanford Dam failed due to overtopping and took at least 30 minutes based on drone photos (France et al. 2022); we estimate one hour, which is a moderate value given other dam failures (Walder and O'Connor 1997).

- V_0: France et al. (2022) states that "Wixom Lake was drained" over the initial breaching. Satellite images suggest that at least some water in the western, Tobacco River half of Wixom Lake was unable to drain due to a knickpoint near the causeway that had divided the two halves and knickpoints #1-4. As such, we estimate the volume released by the failure to be 80% of the original dam capacity. The Sanford dam failed from overtopping and nearly completely drained its reservoir, and so we use its original dam capacity.

2. Values of η from Step One were compared to a range of about 0.6 to 5, with values outside of this range able to use an asymptotic relationship (equations 20a and 20b in Walder and O'Connor 1997; the references to equations 21a and 22b in step 2 of section 6 are believed to be typos). Edenville's η value was 866 (due to the extraordinarily rapid breaching) and Sanford's η value was 7, justifying for both cases the use of equation 20b:

$$Q_p = 1.94 * g^{\frac{1}{2}} * d^{\frac{5}{2}} * \left(\frac{D_c}{d}\right)^{\frac{3}{4}}$$ (Supplemental Eq. 2)

where Q_p is estimated peak discharge (m^3/s) and D_c is original dam height (m).

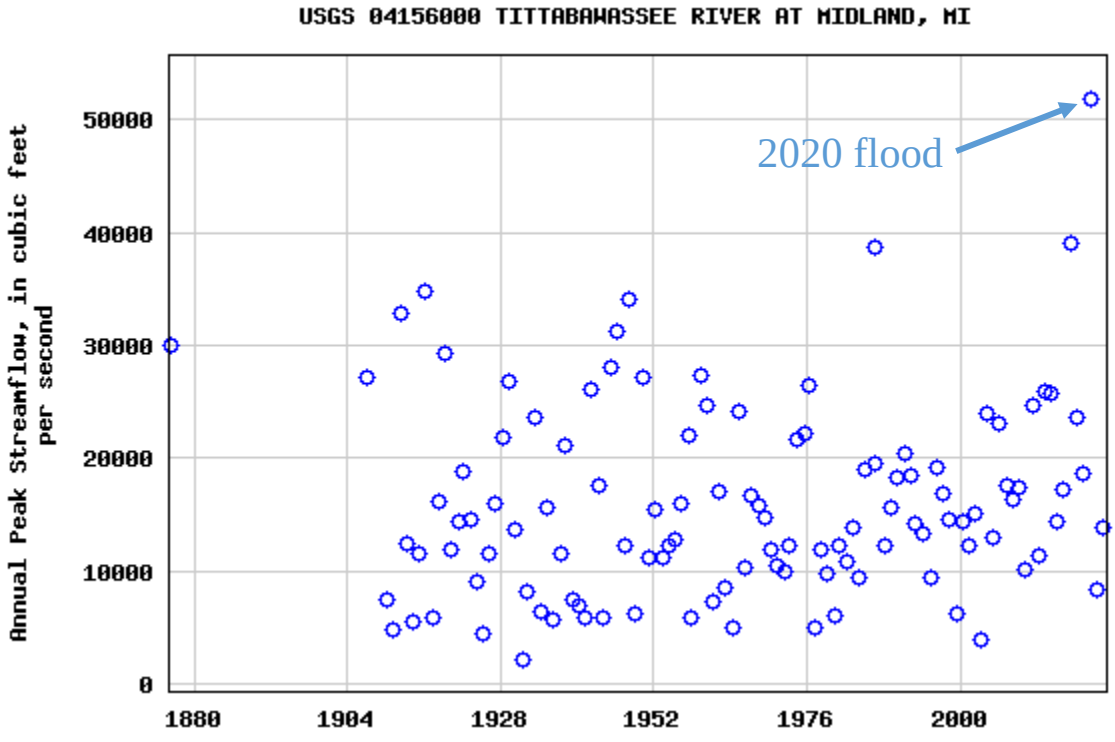
- D_c: engineering reports state that the dam breach was complete/total at Edenville (France et al. 2022), and so $D_c = d$. This number is higher than the original dam height reported in Table 1 of the main manuscript because the failure occurred away from the spillway onto an area of higher elevation. The same proportion was not reported for Sanford, and so we used the pre-failure reported dam height (NID 2022b).

These methods yield estimates of approximately 2,440 m^3/s at Edenville and 1,580 m^3/s. These values seem reasonable; Edenville should be higher than Sanford, considering it was a larger dam with a more rapid failure and no opportunity to buffer any upstream flood

297 waves, as would have occurred leading to Sanford. Both peaks are also higher than, but of
298 a similar magnitude to, the peak discharge observed approximately 20 km downstream at
299 the USGS gauge station (04156000; USGS 2020) at Midland, MI (Supplemental Figure
300 16).
301

Supporting Table 8: Values of parameters used for peak discharge estimation following Walder and O'Connor (1997). Methods and sources for each value are explained in the above text.

Parameter	Symbol	Units	Edenville	Sanford
mean vertical erosion rate of breach	k	m/s	0.183333	0.002389
depth of breach erosion	d	m	11	8.6
breach-formation time	t	s	60	3600
volume of water released	V_0	m^3	65,280,000	17,000,000
acceleration due to gravity	g	m/s^2	9.81	9.81
non-dimensional product of lake volume and erosion rate	η	-	866	6.95
original dam height	D_c	m	11	11
estimated peak discharge	Q_p	m^3/s	2,440	1,590



Supporting Figure 16: Annual peak streamflow for the USGS gauge station 04156000 at Midland, MI. The 2020 flood was the largest on record, but not by an exceptional amount.